AFHRL-TP-89-52

AIR FORCE

HUMAN RESOURCES

AD-A221 674

AIR FORCE OFFICER QUALIFYING TEST (AFOQT): PREDICTORS OF UNDERGRADUATE PILOT TRAINING AND UNDERGRADUATE NAVIGATOR TRAINING SUCCESS

Thomas O. Arth, 1Lt, USAF
Kurt W. Steuck
Christopher T. Sorrentino, 2Lt, USAF
Eugene F. Burke, U.K. Exchange Psychologist

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

May 1990
Interim Technical Report for Period April 1987 - August 1989

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

NOTICE

This paper has been reviewed and is approved for publication.

WILLIAM E. ALLEY, Technical Director
Manpower and Personnel Division

DANIEL L. LEIGHTON, Colonel, USAF
Chief, Manpower and Personnel Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
May 1990

3. REPORT TYPE AND DATES COVERED
Interim - April 1987 to August 1989

4. TITLE AND SUBTITLE
Air Force Officer Qualifying Test (AFOQT): Predictors of Undergraduate Pilot Training and Undergraduate Navigator Training Success

5. FUNDING NUMBERS
PE - 62703F
PR - 7719
TA - 18
WU - 47

6. AUTHOR(S)
Thomas O. Arth
Kurt W. Steuck
Christopher T. Sorrentino
Eugene F. Burke

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Manpower and Personnel Division
Air Force Human Resources Laboratory
Brooks Air Force Base, Texas 78235-5601

8. PERFORMING ORGANIZATION REPORT NUMBER
AFHRL-TP-89-52

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This research investigated the relationship between the Air Force Officer Qualifying Test (AFOQT) Form O and performance (pass/fail) in Undergraduate Pilot Training and Undergraduate Navigator Training. It was found that the several subtests and composites currently being used in the pilot and navigator selection had significant correlations with pilot and navigator training success, respectively. When the correlations were corrected for restriction in range, the correlations increased moderately. Regression analysis revealed that the AFOQT has greater accuracy in predicting success in UPT and UNT when two distinct composites are used than when one combined composite is used. The evaluation of the potential composites against the existing composites revealed that several alternative composites were more effective in predicting pilot and navigator training success than those currently in operational use.

14. SUBJECT TERMS
Air Force Officer Qualifying Test, navigator selection
aptitude tests, pilot selection

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

AFHRL Technical Paper 89-52 May 1990

AIR FORCE OFFICER QUALIFYING TEST (AFOQT): PREDICTORS OF UNDERGRADUATE PILOT TRAINING AND UNDERGRADUATE NAVIGATOR TRAINING SUCCESS

Thomas O. Arth, 1Lt, USAF
Kurt W. Steuck
Christopher T. Sorrentino, 2Lt, USAF
Eugene F. Burke, U.K. Exchange Psychologist

MANPOWER AND PERSONNEL DIVISION
Brooks Air Force Base, Texas 78235-5601

Reviewed by

David E. Brown, Lt Col, USAF
Chief, Officer Selection and Classification Function

Submitted for publication by

Lonnie D. Valentine, Jr.
Chief, Force Acquisition Branch

This publication is primarily a working paper. It is published solely to document work performed.

SUMMARY

The Air Force Officer Qualifying Test (AFOQT) is used to select qualified applicants to attend Air Force Officer Training School (OTS) and the Air Force Reserve Officer Training Corps (AFROTC) program. The AFOQT is also used to select those qualified for pilot training or navigator training.

Four goals were pursued in this effort. The first was to correlate the existing subtests and composites against performance in pilot and navigator training programs. The second was to correct these correlation coefficients for restriction in range to estimate the true relationship between scores and training outcomes. The third goal was to determine whether one "aircrew" composite could be constructed and used to select pilots and navigators. The fourth was to construct and evaluate a variety of alternative composites which could be used for pilot and navigator selection.

Zero order correlations were calculated to validate the existing AFOQT subtests and composites against pilot and navigator training success. A procedure was then used to correct for restriction in range due to the highly selective samples used in the analyses. It was found that the several subtests and composites currently being used in pilot and navigator selection had significant correlations with pilot and navigator training success, respectively. When corrected for restriction in range, the corrrelations increased moderately.

Regression models were used to investigate the possibility of constructing a single aircrew composite for use in selection and classification. Regression analysis revealed that it is not possible to construct a combined composite from the current AFOQT subtests that could be used in selecting both pilots and navigators.

Finally, alternative composites were developed on the basis of their greatest zero order correlations with training pass/fail and their respective standard error of estimate. On evaluation it was revealed that several new composites were more effective in predicting pilot and navigator training success.

PREFACE

This study was completed under Task 771918, Selection and Classification Technologies, which is part of a larger effort in Force Acquisition and Distribution. It was subsumed under work unit 77191847, Development and Validation of Civilian and Non-rated Officer Selection Methodologies. This work unit was established in response to Air Force Regulation 35-8, Air Force Military Personnel Testing System.

The authors would like to thank several members of the Air Force Human Resources Laboratory for their assistance during this project. They include Ms. Doris Black, Mr. Cal Fresne, Mr. James Friemann, SSgt David LeBrun, Mr. John Graves, Sgt Paul Schexnayder, SrA Robert Holyman, Mr. Lewis Walker, and Ms. Suzi Farrell. Thanks are also due to Dr. Malcolm Ree for giving us many hours of consultation on this matter.

TABLE OF CONTENTS

LIST OF TABLES

AIR FORCE OFFICER QUALIFYING TEST (AFOQT):
PREDICTORS OF UNDERGRADUATE PILOT TRAINING AND
UNDERGRADUATE NAVIGATOR TRAINING SUCCESS

I. INTRODUCTION

The Air Force Officer Qualifying Test (AFOQT) is a paper-and-pencil multiple aptitude test battery used to select individuals for Officer Training School (OTS) and Air Force Reserve Officer Training Corps (AFROTC). In general, individuals wishing to enroll in the AFROTC program take the AFOQT prior to their freshman year in college, if applying for a 4-year scholarship, otherwise prior to their junior year in college. Selection is based on the Verbal, Quantitative, and Academic Aptitude scores along with other criteria (e.g., detachment commander rating; see Cowan, Barrett, & Wegner, 1989). Upon selection into the program, cadets enroll in the Professional Officer Course (POC) beginning in the junior year. Successful graduates then enter the Air Force as a second lieutenant. Individuals wishing to enter OTS take the AFOQT during their senior year of college or after having completed the Bachelor's degree. Selection into OTS is based on Academic Aptitude scores and other criteria (Cowan, Barrett, & Wegner, 1990). Successful candidates then complete a 12-week course before being commissioned as a second lieutenant. For both AFROTC and OTS, potential cadets must have a minimum percentile score of 15 on the Verbal composite and 10 on the Quantitative composite.

The AFOQT is also used to select AFROTC and OTS candidates for Undergraduate Pilot Training (UPT) and Undergraduate Navigator Training (UNT). Air Force Regulation 51-4 (1986), provides guidance on the use of the Pilot and Navigator-Technical composites. To be considered for UPT, applicants must score at a minimum of the 25th percentile on the Pilot composite, 10th percentile on Navigator-Technical composite, and a combination of 50 when Pilot and Navigator-Technical are added together. UNT applicants must score at least at the 25th percentile on Navigator-Technical, 10th percentile on Pilot, and 50 when Pilot and Navigator-Technical are added together. Thus, both composites are actually used for selection purposes for both training schools.

The AFOQT Form O is composed of 380 items which form 16 subtests (Rogers, Roach, & Wegner, 1986; Skinner & Ree, 1987). The subtests are aggregated into 5 composites (Table 1). These are: Pilot (P), Navigator-Technical (N-T), Academic Aptitude (AA), Verbal (V), and Quantitative (Q). To compute the composite scores subtest raw scores are added together and converted to percentile scores which are used for selection.

The purpose of this research was fourfold. The first was to validate the existing subtests and composites for performance in UPT and UNT. The approach taken was to correlate AFOQT raw scores with pass/fail in UPT or UNT as the criterion. The second goal was to correct these correlation coefficients for restriction in range. The correction was necessary because the validity coefficients were computed on highly selected samples. This underestimates the true relationship between the test scores and training outcomes. The third goal was to determine whether the existing Navigator-Technical and Pilot composites could be combined into a single "aircrew" composite to be used in the selection of officer applicants into either UPT or UNT. It is possible that a single set of predictors could be used in both sets of training courses. The fourth goal was to construct and evaluate a variety of alternative Pilot and Navigator-Technical composites in the event that one "aircrew" composite is not feasible. A decade ago Hunter and Thompson (1978) identified the best set of subtests for the Pilot composite while Valentine (1977) did so for the Navigator-Technical composite. This research consolidates and updates those efforts.

Table 1. **Construction of AFOQT Form O Composites**

Subtests	Items	Composites P	N-T	AA	V	Q
Verbal Analogies	25	X		X	X	
Arithmetic Reasoning	25		X	X		X
Reading Comprehension	25			X	X	
Data Interpretation	25		X	X		X
Word Knowledge	25			X	X	
Math Knowledge	25		X	X		X
Mechanical Comprehension	20	X	X			
Electrical Maze	20	X	X			
Scale Reading	40	X	X			
Instrument Comprehension	20	X				
Block Counting	20	X	X			
Table Reading	40	X	X			
Aviation Information	20	X				
Rotated Blocks	15		X			
General Science	20		X			
Hidden Figures	15		X			

II. METHOD

Subjects used in the analyses were officers who took the AFOQT Form O, attended OTS, and then attended either UNT (N = 632) or UPT (N = 695). The subjects ranged in age from 19 to 23 years when they took the AFOQT and between 20 and 28 years when they entered training. The subjects entered training between February 1982 and April 1984 with an average delay of 15 months between the date the AFOQT was taken and the date of entering training. The sample consisted of 98.9% males and 1.1% females; 99% were white, .5% Black, and .5% other.

In order to meet the four goals of this research, several correlational and regression analyses were performed. For the first two goals (i.e., validation of existing subtests and composites with UPT and UNT performance) the predictor variables were the 16 subtests that comprise the AFOQT along with the existing 5 composites. Raw subtest and composite scores, expressed as number of correct responses, were used in the analyses. The predictors used in exploring the possibilitity of a single aircrew composite (i.e., the third goal) consisted of the 16 AFOQT subtest scores and variables which represented whether an individual attended UPT or UNT. The predictors used in evaluating alternate Pilot and Navigator-Technical composites (i.e., the fourth goal) consisted of the 16 subtests in a variety of configurations. All analyses which involved AFOQT used scores from the first time an individual took the AFOQT. The criterion for all of the analyses described below was a dichotomous pass/fail variable.

The validities of existing AFOQT subtests and composites were estimated through Pearson product-moment correlations with pass/fail performance in UPT and UNT. Due to prior selection, the variances of subtests and composites are curtailed in training samples. That is, because only higher scoring applicants are selected for entry to rated training, the variances observed in training samples for subtests and composites will be less than if no selection had occurred and all applicants were admitted to rated training (Thorndike, 1949). Due to these curtailed variances, the observed validity coefficients will tend to underestimate the true validities of subtest and composite scores against UPT and UNT training outcomes.

Subtest validities were corrected for restriction in range using the multivariate procedure developed by Mifflin and Verna (1977). Data provided by Skinner and Ree (1987) for AFOQT

applicant scores were used in correcting the subtest validities. Corrected composite validities were then computed using the formula given by Guilford (1950, p. 587) for composite-criterion correlations. The subtest intercorrelations provided by Skinner and Ree (1987) and the corrected subtest validities obtained through the Mifflin and Verna procedure were used to obtain corrected composite validities.

Regression analysis was used to determine whether the same set of tests could predict graduation from both UPT and UNT. A restricted model was constructed and evaluated against a full model. The full model contained a dichotomous group membership variable indicating whether a person attended UPT, a dichotomous group membership variable indicating whether a person attended UNT, and the interactions between the membership variables and the 16 AFOQT subtest scores. Thus, the full model contained 34 predictor variables. The restricted model also contained both dichotomous group membership variables and the 16 AFOQT subtest scores, but not the interactions of the membership variables with the subtest scores. Thus, this restricted model had 18 predictors. The restricted model was compared to the full model and evaluated using an F-ratio.

The question of which combination and weighting of AFOQT subtests would best predict successful completion of UNT and which subtests would best predict success in UPT was addressed next. Several alternative composites were constructed and compared with the existing Pilot and Navigator-Technical composites. $N\text{-}T_a$ and P_a were designed to include those subtests with the greatest uncorrected zero-order correlations. Subtests with positive regression weights were used to form three additional alternative composites for the two rated composites. $N\text{-}T_b$ and P_b included subtests with double, unit, and half weighting. $N\text{-}T_c$ and P_c included the same subtests as $N\text{-}T_b$ and P_b, but with slightly different configurations of double, unit, and half weighting. $N\text{-}T_d$ and P_d contained the same subtests, but all had unit weighting. Subtests in $N\text{-}T_e$ and P_e were chosen based upon the error mean squares computed in a stepwise regression analysis using all 16 subtests. The point at which the least change occurred in the error mean square was identified. All subtests included in the stepwise regression were included in the proposed composites. Composites $N\text{-}T_f$ and P_f were composed of a subset of $N\text{-}T_e$ and P_e with only those subtests having positive regression weights included.

III. RESULTS

The proportion of officers completing UPT in this sample is .79 with a standard deviation of .41. The means and standard deviations of the AFOQT subtests and composites for the pilot sample are presented in Table 2. A comparison of the means with those of the general applicant population (Table A-1) reveals that on all subtests pilots score higher than the general population. The difference tends to be approximately two to three raw score points for every 20 items. The greatest differences are found on the Instrument Comprehension and Aviation Information subtests. On the composites, the difference between the general applicant population and pilots also tends to be about two points for every 20 items with the exception of the Pilot composite. On that composite, the difference is approximately three points for every 20 items. The standard deviations for the two groups are also not identical. For each subtest and composite, the standard deviation for the pilot sample is less than that of the general applicant population indicating that restriction in range had occurred.

The proportion of officers in this sample completing UNT is .86 with a standard deviation of .34. Table 3 contains the means and standard deviations for the AFOQT subtests and composites for the navigator sample. In general, this sample tends to score higher on average than the general applicant population by one to three items for every 20 items. The subtests with the greatest differences, Reading Comprehension and Word Knowledge, are both verbal in

Table 2. **Descriptive Statistics of AFOQT Subtest and Composite Scores for Pilot Sample (N = 695)**

Subtest	Number of items	Mean	SD
Verbal Analogies	22	15.14	3.20
Arithmetic Reasoning	21	13.41	3.53
Reading Comprehension	25	19.21	4.33
Data Interpretation	23	13.78	3.41
Word Knowledge	24	15.83	4.92
Math Knowledge	25	16.12	4.97
Mechanical Comprehension	19	12.50	3.12
Electrical Maze	20	9.34	4.04
Scale Reading	39	24.59	5.10
Instrument Comprehension	20	14.33	4.20
Block Counting	20	13.03	3.53
Table Reading	40	30.52	5.74
Aviation Information	20	14.94	3.86
Rotated Blocks	15	9.17	2.79
General Science	20	10.57	3.22
Hidden Figures	15	10.56	2.54
Composite			
Pilot	200	134.37	17.80
Navigator-Technical	257	163.57	24.91
Academic Aptitude	140	93.49	17.77
Verbal	71	50.18	10.82
Quantitative	69	43.31	9.91

Table 3. **Descriptive Statistics of AFOQT Subtest and Composite Scores for Navigator Sample (N = 632)**

Subtest	Number of items	Mean	SD
Verbal Analogies	22	15.72	3.10
Arithmetic Reasoning	21	13.49	3.59
Reading Comprehension	25	19.79	4.11
Data Interpretation	23	13.48	3.43
Word Knowledge	24	17.13	4.85
Math Knowledge	25	16.32	4.67
Mechanical Comprehension	19	11.80	3.01
Electrical Maze	20	8.88	4.04
Scale Reading	39	23.88	5.13
Instrument Comprehension	20	11.04	4.38
Block Counting	20	12.89	3.70
Table Reading	40	30.00	5.83
Aviation Information	20	10.56	3.98
Rotated Blocks	15	8.88	2.82
General Science	20	10.01	3.24
Hidden Figures	15	10.48	2.46
Composite			
Pilot	200	124.79	18.76
Navigator-Technical	257	160.13	24.52
Academic Aptitude	140	95.93	16.75
Verbal	71	52.64	10.34
Quantitative	69	43.29	9.64

nature. On the composites, navigators score about two raw score points higher per 20 items on four composites, but approximately three points higher per 20 items on the Verbal composite. The pattern for the standard deviations for the navigators is similar to that of the pilots. In all cases the standard deviations are lower than those of the general applicant population, again indicating that restriction in range had occurred.

The subtest intercorrelation matrices for the pilot and navigator samples are presented in Tables 4 and 5, respectively. One salient feature of these matrices is that the correlations between subtests for the two samples are lower than the corresponding correlations in the general applicant population (Table A-2). Again the restriction in scores is evident.

The composite intercorrelation matrices for the two samples are presented in Tables 6 and 7, respectively. Caution is needed when interpreting these tables, because the correlations are spuriously inflated due to common subtests appearing across composites (see Table 1). As with the subtest intercorrelations, composite intercorrelations for the pilots and navigators are generally lower than those of the general population (Table A-3).

Table 8 presents the uncorrected and corrected zero order correlations of the 16 AFOQT subtests and five existing composites with success in UNT and UPT. Overall, the subtests that are currently being used in the Pilot and Navigator-Technical composites had significant correlations with success in UPT and UNT, respectively. One interesting result of the correction for range restriction procedure is that an uncorrected correlation that is negative in direction may change to a positive corrected correlation. This is because an entire system of variances and covariances is used in correcting feature correlations.

In the UNT sample, five subtests were found to yield corrected correlations with UNT success equal to or greater than .20, those being Arithmetic Reasoning, Math Knowledge, Scale Reading, Block Counting, and Rotated Blocks. All five subtests are included in the current Navigator-Technical composite. However, the General Science subtest (a constituent of the current Navigator-Technical composite) was not found to yield a significant zero-order correlation with UNT success.

Overall, the subtests in the Navigator-Technical composite with significant correlations can be described as being visual-spatial or quantitative in nature. For example, among the subtests with the highest correlations were Scale Reading and Arithmetic Reasoning. No subtest having a large verbal component (e.g., Verbal Analogies, Reading Comprehension) had a significant correlation with success in UNT.

In the UNT sample, three of the five AFOQT composites had significant uncorrected validities with UNT outcome (Pilot, Navigator-Technical, and Quantitative). When corrected for restriction in range, three composites had correlations with UNT performance that were greater than .20 (Pilot, Navigator-Technical, and Quantitative).

In the UPT sample, four of the eight subtests in the existing Pilot composite had significant uncorrected correlations with success in UPT. These were Mechanical Comprehension, Scale Reading, Instrument Comprehension, and Aviation Information. Four subtests comprising the Pilot composite failed to reach significance (Verbal Analogies, Electrical Maze, Block Counting, and Table Reading). In addition, three subtests that are not in the current Pilot composite were significantly related to the criterion. They were Data Interpretation, Word Knowledge, and Rotated Blocks. The Word Knowledge subtest correlation was in a negative direction, but this is most likely due to range restriction.

Table 4. **Intercorrelation Matrix of AFOQT Subtests for Pilot Sample (N = 695)**

Subtest	VA	AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS	HF
Verbal Analogies	1.00															
Arithmetic Reasoning	.38	1.00														
Reading Comprehension	.58	.38	1.00													
Data Interpretation	.31	.57	.35	1.00												
Word Knowledge	.58	.28	.69	.24	1.00											
Math Knowledge	.40	.56	.40	.47	.32	1.00										
Mechanical Comprehension	.28	.27	.30	.25	.23	.26	1.00									
Electrical Maze	.11	.19	.14	.23	.05	.25	.30	1.00								
Scale Reading	.20	.49	.24	.50	.15	.44	.18	.25	1.00							
Instrument Comprehension	.08	.15	.02	.23	.02	.10	.30	.25	.23	1.00						
Block Counting	.26	.31	.19	.31	.12	.35	.27	.31	.37	.22	1.00					
Table Reading	.12	.27	.09	.30	.01	.27	.04	.14	.36	.14	.37	1.00				
Aviation Information	-.05	.00	-.08	.06	-.09	-.15	.22	-.05	.00	.47	-.10	-.09	1.00			
Rotated Blocks	.27	.25	.19	.28	.18	.29	.39	.28	.29	.27	.35	.14	.02	1.00		
General Science	.35	.24	.39	.22	.35	.34	.42	.19	.11	.21	.14	.04	.16	.30	1.00	
Hidden Figures	.22	.14	.15	.16	.12	.24	.24	.18	.24	.14	.27	.16	-.05	.32	.15	1.00

Table 5. **Intercorrelation Matrix of AFOQT Subtests for Navigator Sample (N = 632)**

Subtest	VA	AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS	HF
Verbal Analogies	1.00															
Arithmetic Reasoning	.31	1.00														
Reading Comprehension	.54	.32	1.00													
Data Interpretation	.31	.53	.35	1.00												
Word Knowledge	.57	.25	.66	.28	1.00											
Math Knowledge	.29	.57	.30	.45	.23	1.00										
Mechanical Comprehension	.28	.18	.23	.16	.16	.21	1.00									
Electrical Maze	.17	.21	.05	.21	.00	.21	.30	1.00								
Scale Reading	.17	.54	.15	.45	.09	.43	.18	.29	1.00							
Instrument Comprehension	.17	.11	.09	.22	.08	.14	.32	.26	.24	1.00						
Block Counting	.27	.24	.17	.21	.12	.24	.23	.30	.33	.26	1.00					
Table Reading	.09	.22	.10	.26	.07	.23	.05	.17	.34	.15	.39	1.00				
Aviation Information	.05	.02	.06	.06	.11	.03	.38	.14	.08	.46	.04	.01	1.00			
Rotated Blocks	.22	.23	.10	.21	.10	.26	.41	.33	.35	.29	.40	.20	.18	1.00		
General Science	.33	.21	.37	.22	.36	.35	.46	.19	.16	.31	.13	.02	.37	.22	1.00	
Hidden Figures	.20	.25	.16	.17	.13	.20	.19	.19	.29	.17	.24	.17	.11	.30	.17	1.00

Table 6. **Correlation Matrix of AFOQT Composites for Pilot Sample (N = 695)**

Composite	Pilot	Navigator-Technical	Academic Aptitude	Verbal	Quantitative
Pilot	1.00				
Navigator-Technical	.86	1.00			
Academic Aptitude	.53	.73	1.00		
Verbal	.34	.43	.87	1.00	
Quantitative	.57	.83	.84	.47	1.00

Note. The correlations are inflated, since the composites have several subtests in common.

Table 7. **Correlation Matrix of AFOQT Composites for Navigator Sample (N = 632)**

Composite	Pilot	Navigator-Technical	Academic Aptitude	Verbal	Quantitative
Pilot	1.00				
Navigator-Technical	.87	1.00			
Academic Aptitude	.50	.69	1.00		
Verbal	.35	.38	.85	1.00	
Quantitative	.50	.79	.83	.41	1.00

Note. The correlations are inflated, since the composites have several subtests in common.

The subtests that yielded the highest corrected validities with success in UPT tended to be visual-spatial or highly related to specific flying information. The two subtests with the highest correlations are Instrument Comprehension and Aviation Information. Others that had large correlations included Rotated Blocks, Scale Reading, and Block Counting, all visual-spatial type tasks. Verbal subtests were found to have low validities. This may be due to requiring OTS applicants to have a college degree which curtails the variance in tests in the AFOQT.

Of the uncorrected composite validities in the UPT sample, only the Pilot and the Navigator-Technical composites were positively and significantly related to UPT outcome. However, when the attenuation correction procedures were applied all five AFOQT composites were positively related to UPT outcome, although the Verbal composite showed a low corrected validity with UPT success.

In examining the question of whether a common aircrew composite could be used in place of separate Pilot and Navigator-Technical composites F-tests were used to compare a full regression model with a restricted model. This test revealed a significant difference between the restricted (R = .31) and full (R = .37) model indicating that there was a significant loss in the ability to predict success in UPT and UNT when a common regression equation is used, F (16, 1293) = 3.65, p < .001. This means that it is better to have separate composites for selection into UNT and UPT.

The question of which combination and weighting of AFOQT subtests would best predict successful completion of UNT and UPT was addressed next. Table 9 shows the alternative composites with their resulting uncorrected correlations and standard errors of estimate (SEE). For the Navigator-Technical composite, zero-order correlations ranged from .15 to .23. The highest correlations and lowest SEEs were found in N-$T_{b, c, d,}$ and $_f$. Based on a principle of parsimony, N-T_f is the most desirable, because it has the fewest number of subtests forming the composite. Nonetheless, the most improved Navigator-Technical composite would account for only a small increase in predictive accuracy over the existing Navigator-Technical composite.

Table 8. **Correlations of AFOQT Subtests and Composites with UNT (N = 632) and UPT (N = 695) Pass/Fail Criteria[a]**

	UNT		UPT	
	Uncorrected	Corrected	Uncorrected	Corrected
Subtest				
Verbal Analogies	-.06	.04	-.05	.09
Arithmetic Reasoning	.18****	.23	.04	.16
Reading Comprehension	-.04	.05	-.04	.12
Data Interpretation	.09*	.17	.09*	.19
Word Knowledge	-.06	.01	-.11***	.04
Math Knowledge	.14****	.21	-.01	.15
Mechanical Comprehension	.07	.17	.11***	.21
Electrical Maze	.09*	.16	.05	.15
Scale Reading	.19****	.25	.09*	.21
Instrument Comprehension	.06	.16	.34****	.38
Block Counting	.15****	.24	.04	.19
Table Reading	.13****	.19	.06	.16
Aviation Information	.06	.13	.27****	.30
Rotated Blocks	.14****	.20	.11***	.23
General Science	-.01	.07	.02	.13
Hidden Figures	.11**	.17	.05	.16
Multiple R	.255		.395	
Composites				
Pilot	.17****	.25	.21****	.30
Navigator-Technical	.21****	.27	.10**	.24
Academic Aptitude	.06	.14	-.03	.15
Verbal	-.06	.04	-.08*	.09
Quantitative	.17****	.23	.04	.19

*p < .05.

**p < .01.

***p < .005.

****p < .001.

[a]To date, there is no well accepted significant test for correlations corrected for attenuation.

Greater increases in observed validities were noted in the alternative Pilot composites. While the existing Pilot composite's correlation with the UPT criterion was .21, the P_a and P_b composites' correlations were .30. P_a and P_b composites also had lower SEEs than the existing Pilot composite. Thus, the existing Pilot composite has slightly less predictive accuracy than the alternative composites P_a and P_b. Nonetheless, the increases in the ability to predict success in UPT and UNT found here are not sufficient to warrant recommended changes in the composites at this time.

Table 9. **Uncorrected Correlations Between AFOQT Predictors and Pass/Fail for Existing and Alternative Navigator-Technical (N-T) and Pilot (P) Composites**

Composite		Subtest content	R	SEE
Existing N-T	=	AR + DI + MK + MC + EM + SR + BC + TR + RB + GS + HF	.21	.336
N-T_a	=	AR + DI + MK + SR + BC + TR + RB + HF	.22	.334
N-T_b	=	2(AR + BC) + MK + MC + SR + AI + RB + HF + .5(TR)	.23	.334
N-T_c	=	2(AR) + MK + BC + AI + RB + HF + .5(MC + SR + TR)	.23	.334
N-T_d	=	AR + MK + MC + SR + BC + TR + AI + RB + HF	.23	.334
N-T_e	=	VA + AR + WK + MK + SR + BC + AI + RB + GS + HF	.15	.339
N-T_f	=	AR + MK + SR + BC + AI + RB + HF	.23	.334
Existing P	=	VA + MC + EM + SR + IC + BC + TR + AI	.21	.398
P_a	=	MC + SR + IC + AI + RB	.30	.387
P_b	=	2(IC + AI + RB) + RC + .5(AR + DI + MC + SR + TR + HF)	.30	.387
P_c	=	2(IC) + RC + MC + AI + RB + .5(AR + DI + SR + TR + HF)	.28	.391
P_d	=	AR + RC + DI + MC + SR + IC + TR + AI + RB + HF	.21	.398
P_e	=	RC + WK + EM + IC + AI + RB + GS	.15	.402
P_f	=	RC + IC + AI + RB	.28	.390

Note. N-$T_{a\text{-}f}$ = Navigator-Technical Alternative Composites, $P_{a\text{-}f}$ = Pilot Alternative Composites, VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Math Knowledge, MC = Mechanical Comprehension, EM = Electrical Maze, SR = Scale Reading, IC = Instrument Comprehension, BC = Block Counting, TR = Table Reading, AI = Aviation Information, RB = Rotated Blocks, GS = General Science, and HF = Hidden Figures.

IV. SUMMARY

The findings in this study support the use of specialized aircrew composites to select pilots and navigators as is current practice. A single equation that predicted success in both UPT and UNT was found to be significantly different than an equation which predicted success in UPT and UNT separately. This means that predicting success in UPT and UNT is best handled by two distinct composites. In exploring the validities of subtests and composites, several subtests proved to be more effective than others in predicting success in UNT and UPT. Some of the subtests in the existing composites were not useful in predicting pass/fail in UNT or UPT. Conversely, some subtests not included in the existing composites were found to have some predictive validity for navigator and pilot training. Furthermore, from the preliminary construction of alternative composites, it appears that while new configurations of subtests might result in composites which have equivalent predictive validity using fewer subtests, none of the alternatives provided a practical increase in validity.

It is interesting to note however that an optimal equation (using both positive and negative weights) had an uncorrected validity of .395 in predicting UPT success as compared to a corrected validity of .30 for the current pilot composite. An uncorrected validity of .40 probably represents an "upper bound" estimate of how much improvement potential exists in this area. Future efforts will need to explore more creative strategies for composite construction if this potential is to be realized.

REFERENCES

Air Force Regulation 35-8. (1983, April). *Air Force Military Personnel Testing System.* Randolph AFB, TX: AFMPC/DPMYO.

Air Force Regulation 51-4. (1986, August). *Applications procedures for Undergraduate Flying Training.* Randolph AFB, TX: AFMPC/DPMR.

Cowan, D.K., Barrett, L.E., & Wegner, T.G. (1989). *Air Force Reserve Officer Training Corps Selection System Validation* (AFHRL-TR-88-54, AD-A218 494). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Cowan, D.K., Barrett, L.E., & Wegner, T.G. (1990). *Air Force Officer Training School Selection System Validation* (AFHRL-TR-89-65). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Guilford, J.P (1950). *Fundamental statistics in psychology and education* (2nd. ed.). New York: McGraw-Hill.

Hunter, D.R., & Thompson, N.A. (1978). *Pilot selection system development* (AHFRL-TR-78-33, AD-A058 418). Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

Mifflin, T.L., & Verna, S.M. (1977). *A method to correct correlation coefficients for the effects of multiple curtailment* (CRC-336). Arlington, VA: Marine Corps Operations Analysis Group, Center for Naval Analyses.

Rogers, D.L., Roach, B.W., & Wegner, T.G. (1986). *Air Force Officer Qualifying Test Form O: Development and standardization* (AFHRL-TR-86-24, AD-A172 037). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Skinner, M.J., & Ree, M.J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form O* (AFHRL-TR-86-68, AD-A184 975). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Thorndike, R.L. (1949). *Personnel selection.* New York: Wiley.

Valentine, L.D., Jr. (1977). *Navigator-observer selection research: Development of a new Air Force Officer Qualifying Test navigator-technical Composite* (AFHRL-TR-77-36, AD-A042 689). Brooks AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

APPENDIX A: DESCRIPTION OF SAMPLE USED IN RANGE RESTRICTION CORRECTION

Table A-1. Descriptive Statistics of AFOQT and Composite Subtest Scores (N = 3000)

	Number of items	Mean	SD	Skew	Kurtosis	Reliability[a]
Subtest						
Verbal Analogies	22	13.36	4.23	-.39	-.40	.80
Arithmetic Reasoning	21	11.00	4.40	.07	-.66	.81
Reading Comprehension	25	15.83	5.93	-.30	-.93	.88
Data Interpretation	23	11.15	3.93	.18	-.36	.71
Word Knowledge	24	13.28	5.83	.08	-.99	.88
Math Knowledge	25	14.48	6.04	-.04	-1.07	.88
Mechanical Comprehension	19	9.78	3.65	.01	-.58	.71
Electrical Maze	20	7.68	4.22	.75	.24	.81
Scale Reading	39	20.07	6.73	-.03	-.37	.84
Instrument Comprehension	20	8.82	4.76	.36	-.69	.84
Block Counting	20	10.62	4.39	-.08	-.58	.83
Table Reading	40	26.46	7.35	-.50	.50	.92
Aviation Information	20	8.65	4.08	.56	-.16	.77
Rotated Blocks	15	7.59	3.36	-.06	-.76	.77
General Science	20	8.54	3.66	.42	-.29	.70
Hidden Figures	15	9.60	2.76	-.32	.03	.69
Composite						
Pilot	200	105.44	27.84	-.17	-.21	
Navigator-Technical	257	136.96	36.58	-.12	-.36	
Academic Aptitude	140	79.10	24.56	-.12	-.71	
Verbal	71	42.47	14.51	-.19	-.84	
Quantitative	69	36.63	12.66	.07	-.79	

[a]Estimated by coefficient alpha.

Table A-2. **Intercorrelation Matrix of AFOQT Subtests (N = 3000)**

Subtest	VA	AR	RC	DI	WK	MK	MC	EM	SR	IC	BC	TR	AI	RB	GS	HF
Verbal Analogies	1.00															
Arithmetic Reasoning	.58	1.00														
Reading Comprehension	.73	.58	1.00													
Data Interpretation	.53	.67	.55	1.00												
Word Knowledge	.68	.46	.77	.46	1.00											
Math Knowledge	.55	.71	.51	.60	.40	1.00										
Mechanical Comprehension	.48	.51	.46	.46	.40	.48	1.00									
Electrical Maze	.27	.37	.23	.38	.17	.40	.44	1.00								
Scale Reading	.48	.66	.45	.62	.37	.60	.48	.45	1.00							
Instrument Comprehension	.34	.41	.33	.43	.28	.39	.49	.44	.49	1.00						
Block Counting	.45	.53	.40	.51	.32	.49	.50	.47	.61	.49	1.00					
Table Reading	.34	.44	.35	.47	.27	.44	.30	.31	.56	.34	.51	1.00				
Aviation Information	.30	.31	.34	.34	.32	.25	.50	.29	.33	.56	.31	.21	1.00			
Rotated Blocks	.43	.47	.35	.42	.29	.49	.54	.42	.49	.46	.55	.34	.34	1.00		
General Science	.51	.49	.55	.44	.51	.52	.57	.34	.41	.41	.37	.25	.46	.40	1.00	
Hidden Figures	.40	.40	.36	.39	.31	.40	.39	.34	.47	.36	.45	.36	.27	.42	.34	1.00

Table A-3. Intercorrelations Among Composites (N = 3000)

	Pilot	Navigator-Technical	Academic Aptitude	Verbal	Quantitative
Pilot	1.00				
Navigator-Technical	.95	1.00			
Academic Aptitude	.75	.83	1.00		
Verbal	.62	.63	.92	1.00	
Quantitative	.76	.89	.89	.63	1.00

Note. The correlations are inflated, since the composites have several subtests in common.